

# Cooperación de perceptores multicapa y regresión contraída para la estimación de iluminantes<sup>1</sup>

EDWIN OCTAVIO GÓMEZ G.<sup>2</sup>  
HUMBERTO LOAIZA<sup>3</sup>  
EDUARDO CAICEDO<sup>4</sup>



## Resumen

El presente artículo revela un enfoque original para la estimación del iluminante que aprovecha la cooperación entre dos máquinas de aprendizaje: Redes neuronales artificiales y Regresión contraída, tomando las ventajas de cada una. En el documento se exponen los antecedentes teóricos y prácticos de la investigación, la metodología que permitió la implementación y la experimentación del enfoque cooperativo. Los resultados obtenidos en la exploración de condiciones y parámetros generales demuestran las ventajas del enfoque.

<sup>1</sup> Este trabajo forma parte de la tesis doctoral de Edwin Octavio Gómez González en constancia del color computacional, becario de Colciencias en el programa de apoyo a los doctorados nacionales.

Los autores pertenecen al grupo de investigación Percepción y Sistemas Inteligentes de la Escuela de Ingeniería Eléctrica y Electrónica de la Facultad de Ingeniería de la Universidad del Valle.

<sup>2</sup> Ingeniero Electrónico de la Universidad Nacional de Colombia. Estudiante de doctorado en Ingeniería, becario de Colciencias, edwinmez@univalle.edu.co

<sup>3</sup> Doctor en Robótica de la Université d'Evry, Francia. Profesor de la Universidad del Valle, hloaiza@univalle.edu.co

<sup>4</sup> Doctor en Informática Industrial de la Universidad Politécnica de Madrid. Profesor de la Universidad del Valle, ecaicedo@univalle.edu.co

Fecha de recepción: 10/06/07 Fecha de aprobación: 18/10/07

**Palabras clave**

Estimación del Iluminante, Constancia del Color Computacional, Redes Neuronales Artificiales, Máquinas de aprendizaje, Regresión contraída, Aplicaciones.

**Abstract**

In this paper we introduce an original cooperative approach for illuminant estimation; it takes advantage from two learning machines: Artificial neural networks and Ridge Regression. The paper exposes theoretical and practical background and methodology for implementation and experiments. Results about the exploration of general conditions and parameters show the advantages for our approach.

**Key words**

Illuminant Estimation, Computational Color Constancy, Artificial Neural Networks, Learning Machines, Regresión contraída, Applications.

**1. Introducción**

La constancia del color computacional puede ser entendida desde el tratamiento digital de imágenes bio-inspirado, como la emulación de la habilidad de algunos primates para percibir como constante la apariencia de un determinado color, a pesar de que haya cambios en los iluminantes y su aplicación puede resultar útil en múltiples tareas que incluyan el tratamiento digital de imágenes de color (Gómez, Loaiza y Caicedo, 2006a).

Muchos investigadores han enfrentado esta tarea llevando a cabo dos pasos básicos, primero se

estima el iluminante de la imagen y luego se aplica la corrección a la imagen usando el llamado modelo diagonal (Barnard, Cardei y Funt, 2002).

La estimación del iluminante puede hacerse de múltiples formas: con asunciones fuertes, como la de que hay una superficie blanca en la imagen (escalar por el máximo) o que el promedio de la luz que llega de la imagen es acromática (Mundo Gris de Buchsbaum, 1980), o haciendo asunciones más débiles como que la norma de Mikowski de luz que sale de la imagen es una constante (Finlayson y Trezzi, 2004) o que la derivada en un espacio espectral opoente de la luz que sale de la imagen es una constante (Van de Weijer y Gevers, 2005) o usando métodos estadísticos como la divergencia KL (Rosenberg, Herbert y Thrun) o usando máquinas de aprendizaje (Argawal, 2005).

Entre las máquinas de aprendizaje usadas están las redes neuronales artificiales RNA (Funt, Cardei y Barnard, 1996), las máquinas de soporte vectorial SVM (Funt y Xiong, 2004) y la regresión contraída RR (Argawal, 2005).

Al interior del grupo de Percepción y Sistemas Inteligentes de la Universidad del Valle se planteó la tesis “Aportes a la constancia del color computacional usando máquinas de aprendizaje” y se usó como punto de partida la implementación existente con redes neuronales artificiales (Cardei, 2000).

La propuesta inicial de la investigación se hizo usando histogramas extraídos del plano de cromaticidad rg y se incorporó el hecho de que otros algoritmos habían mostrado mejoras cuando se pasaba de considerar el plano de cromaticidad rg, para considerar un espacio espectral tridimensional (Barnard, Martin y Funt, 2000; Funt y Xiong, 2004).

Algunas pruebas preliminares evidenciaron que la misma red neuronal produce diferentes resultados cada vez que se entrena, por esto se deben realizar varias repeticiones para extraer estadísticas representativas (Gómez, Loaiza y Caicedo, 2006b). Igualmente, debido a que los tamaños de las bases de entrenamiento, validación y prueba son grandes y que usualmente se usan cientos de neuronas en la capa de entrada, cada entrenamiento en Matlab puede tardar algunas horas (dependiendo de los parámetros elegidos).

Si se consideran los dos aspectos expuestos en el párrafo anterior, la exploración de parámetros se hace lenta porque para cada combinación de parámetros se debe invertir mucho tiempo, así que se aprovecharon los resultados de Argawal (2005), que favorecen a la regresión contraída frente a los perceptrones multicapa (MLP) en la tarea de estimación del iluminante, en cuanto a velocidad y facilidad de implementación, estabilidad de los resultados y desempeño.

Por lo anterior, los autores de la presente investigación proponen una metodología cooperativa que aproveche la velocidad y facilidad de implementación y la estabilidad de los resultados de la regresión contraída para llevar a cabo una exploración inicial sobre el espacio de parámetros generales del problema, para elegir aquellos que ofrezcan mejores resultados, antes de intentar el entrenamiento con un grupo de perceptrones multicapa, en una prueba que tomaría más tiempo, ahorrando tiempo de experimentación y de cómputo.

## 2. Antecedentes

En esta sección se presentan los antecedentes teóricos y prácticos de esta investigación: en primer lugar, se expone la teoría de la regresión

contraída siguiendo a Argawal (2005), para que el lector se familiarice con la técnica. En segundo lugar, se parte de la comparación hecha por el mismo autor entre algunas máquinas de aprendizaje, para introducir la propuesta de un enfoque cooperativo.

### 2.1 La regresión contraída

Según Argawal (2005) la regresión contraída se puede entender como una regresión lineal múltiple con un factor de normalización.

En la regresión lineal se asume que el modelo que describe la relación entre las variables de entrada ( $X$ ) y de salida ( $Y$ ) es lineal y se llama múltiple porque tiene múltiples entradas:

$$Y = f(x) = w_0 + \sum_{j=1}^d w_j x_j = XW \quad (1)$$

Donde los distintos  $w_j$ , que conforman el vector  $W$ , son los llamados coeficientes del modelo. La salida estimada del modelo, se considera la salida real afectada por ruido  $\eta$  de media cero y varianza  $\sigma^2$ .

$$\hat{Y} = Y + \eta \quad (2)$$

Para hallar los coeficientes del modelo se trata de minimizar la diferencia entre la salida real y la salida estimada.

$$\begin{aligned} S &= \sum_{i=1}^n (\eta_i)^2 = \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \\ &= \sum_{i=1}^n \left( \hat{Y}_i - w_0 + \sum_{j=1}^d w_j x_{ij} \right)^2 \\ &= (\hat{Y} - XW)^T (\hat{Y} - XW) \end{aligned} \quad (3)$$

Donde

$$\hat{Y} = (\hat{y}_1, \hat{y}_2, \hat{y}_3, \dots, \hat{y}_n) \quad (4)$$

$$X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nd} \end{pmatrix} \quad (5)$$

Para llevar a cabo la minimización se deriva la ecuación (3) y se iguala a cero.

$$\frac{\partial \eta}{\partial W} = -2X^T (\hat{Y} - XW) = 0 \quad (6)$$

Si  $X$  es no singular,  $(X^T X)^{-1}$ , existe y resolviendo para  $W$ , se obtienen los valores de los coeficientes que minimizan el error:

$$W = (X^T X)^{-1} X^T \hat{Y} \quad (7)$$

Ahora para la regresión contraída, se incluye un factor de normalización que hace más grande el error si los valores de los coeficientes del modelo son muy grandes, equivalente a la Normalización de orden cero de Tikunov.

$$\eta^2 = (\hat{Y}_i - Y_i)^2 + \lambda \sum_{j=1}^d W_j^2 \quad (8)$$

Desarrollando de la misma forma que se hizo en la ecuación (3):

$$\begin{aligned} S_{\text{Contraída}} &= \sum_{i=1}^n (\eta)^2 \\ &= \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 + \lambda \sum_{j=1}^d W_j^2 \\ &= \sum_{i=1}^n \left( \hat{Y}_i - w_0 + \sum_{j=1}^d w_j x_{ij} \right)^2 \\ &\quad + \lambda \sum_{j=1}^d W_j^2 \\ &= (\hat{Y} - XW)^T (\hat{Y} - XW) + \lambda W^T W \end{aligned} \quad (9)$$

Buscando ahora el valor de  $W$  que minimiza la expresión (8), se deriva y se iguala a cero como en (6):

$$\begin{aligned} \frac{\partial S_{\text{Contraída}}}{\partial W} &= \frac{\partial \{ (\hat{Y} - XW)^T (\hat{Y} - XW) + \lambda W^T W \}}{\partial W} \\ \frac{\partial S_{\text{Contraída}}}{\partial W} &= -2X^T \hat{Y} + 2(X^T X + \lambda I)W = 0 \end{aligned} \quad (10)$$

$$W_{\text{Contraída}} = (X^T X + \lambda I)^{-1} X^T \hat{Y} \quad (11)$$

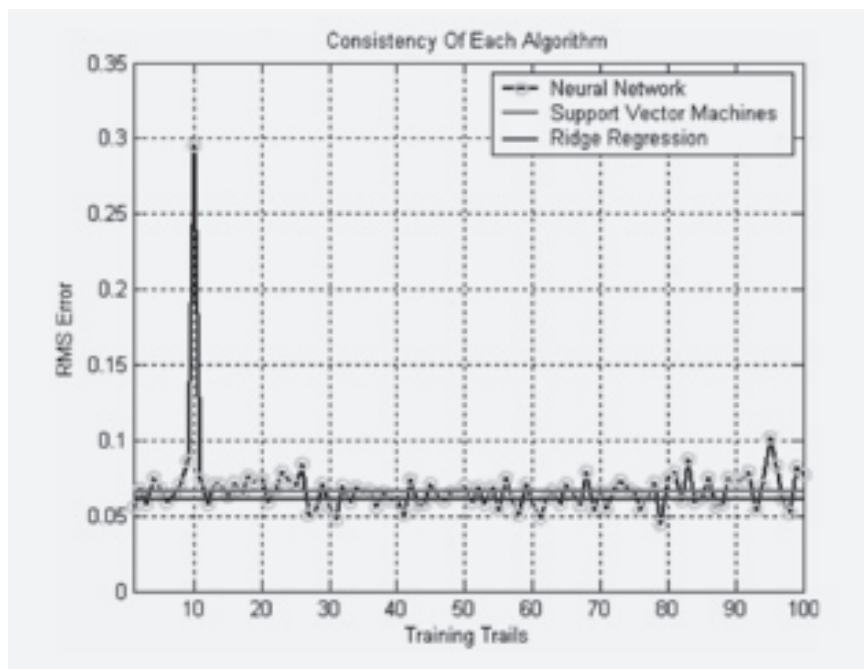
Ahora si se quiere calcular la salida, pues es similar a la ecuación (1):

$$\hat{Y} = XW_{\text{Contraída}} \quad (12)$$

## 2.2 La regresión contraída y el perceptrón multicapa en la estimación del iluminante

Argawal (2005) comparó en su trabajo la regresión contraída, las máquinas de soporte vectorial y el perceptrón multicapa, en la tarea de la estimación de iluminantes. Para hacerlo, entregó a todas las máquinas de aprendizaje el mismo tipo de entradas, de acuerdo con un enfoque propuesto por Funt, Cardei y Barnard (1986) para redes neuronales.

Los resultados de sus experimentos pueden apreciarse en gráficas como la de la Figura 1 y le permitieron concluir que las máquinas de soporte vectorial y la regresión contraída ofrecían respuestas estables para un mismo set de entrenamiento, mientras las redes neuronales daban respuestas diferentes en cada entrenamiento.



**Figura 1.** Errores de predicción obtenidos tras 100 repeticiones de entrenamiento de redes neuronales, máquinas de soporte vectorial y regresión contraída (Fuente: Argawal, 2005).

Asimismo, la descripción que se hizo en 0 pone en evidencia que la implementación de la regresión contraída es más simple y tiene un menor costo computacional que la implementación de una red neuro-

nal o de una regresión por soporte vectorial.

En la presente investigación se ha propuesto realizar algunos aportes a la constancia del color computacional usando máquinas de aprendizaje y para ello se han implementado redes neuronales artificiales, y se ha decidido evaluar la regresión contraída para comparar sus resultados con los de las redes neuronales.

Las pruebas iniciales arrojaron resultados alentadores, porque la estimación del iluminante mediante regresión contraída fue efectivamente más fácil y de menor costo computacional que las redes neuronales.

Lo anterior permitió plantear un método cooperativo y no competitivo, que explorara algunos parámetros mediante la regresión contraída antes de lanzar un entrenamiento con una batería de perceptrones multicapa.

### 3. Metodología

En esta sección se presentan de forma ordenada los procedimientos que permitieron llevar a cabo la experimentación: se explica la síntesis de imágenes y la generación de histogramas a partir de imágenes reales, se detallan las implementaciones de la regresión contraída y el perceptrón multicapa, se expone el enfoque cooperativo propuesto y se definen las medidas de error.

#### 3.1 Generación de imágenes sintéticas

El primer paso para experimentar con constancia del color consiste en generar imágenes sintéticas que no poseen ruido y de las que se conocen con precisión las superficies presentes y los iluminantes usados (Barnard, Cardei y Funt, 2002).

Dichas imágenes se crearon a partir de un grupo de 87 espectros

de iluminantes  $e(\lambda)$  medidos, 1995 reflectancias  $s(\lambda)$  medidas o recopiladas y las funciones de respuesta de la cámara Sony DXC-930,  $R(\lambda)$  estimadas por el SFU Vision Lab (Barnard *et al.*, 2002a), usando una versión discreta (reemplazar integrales por sumatorias) de la ecuación clásica de la tricromancia (13).

$$r_j^i = \int_{\lambda_{\min}}^{\lambda_{\max}} R_j^i(\lambda) c_j(\lambda) d\lambda = \int_{\lambda_{\min}}^{\lambda_{\max}} R_j^i(\lambda) e(\lambda) s_j(\lambda) d\lambda \quad (13)$$

Inicialmente se llevó a cabo una normalización de los iluminantes, dividiendo cada espectro entre la sumatoria de los valores para todas sus longitudes de onda para comparar las formas de los espectros y no su magnitud, este tipo de normalización dio paso a otras propuestas, como se verá en la parte de resultados.

Se generaron 600 escenas sintéticas por cada iluminante para un total de 52.200 escenas sintéticas; cada escena constaba de entre 3 y 80 superficies elegidas por permutación entre las disponibles y el número de reflectancias presentes en cada escena se escogió de manera aleatoria. El número de reflectancias disponibles varió a partir de la colección hecha por Barnard *et al.* 2002a, que tenía 1.995 reflectancias.

En un llamado set inicial sólo se consideraron 251 reflectancias, tratando de aproximarse a lo expuesto por Cardei (2000) en la página 98 de su tesis (260 reflectancias).

Otros experimentos se hicieron con el total de las 1.995 reflectancias, a partir de estas reflectancias disponibles se eligieron, de manera aleatoria, las reflectancias de cada escena sintética.

Se debe aclarar que acorde con la literatura (Cardei, 2000), el término imágenes sintéticas se refiere al



sentido físico que tienen los vectores generados, ya que aunque no se generaron imágenes propiamente, cada vector representa todos los valores tricromáticos que aparecen en una escena sintética.

La base de datos total se dividió en tres bases de tamaños iguales, para construir las bases de entrenamiento, validación y prueba (200 escenas por iluminante - 17.400 escenas en total) de la red neuronal.

Siguiendo la línea de trabajo propuesta por Funt, Cardei y Barnard (1996), cada escena se representó por un código binario que indicaba la presencia (1) o ausencia (0) de ciertos colores, pero como se propuso al inicio de la investigación, se usó un histograma tridimensional del espacio espectral en el cual se analizaba la imagen.

La distribución se hizo mediante una retícula cartesiana con cubos de iguales dimensiones, de tal forma que el tamaño de la arista,  $a$ , determina el número  $N$  de compartimentos del histograma. (14)

$$N = \frac{1}{a^3} \quad (14)$$

Análogo al proceso de poda descrito por Cardei (2000), se despreciaron aquellos componentes del histograma que no se activaron para ninguna superficie de la base de datos.

De esta manera, cada escena sintética quedó representada por un vector binario de dimensiones  $n \times 1$ , donde  $n < N$  y  $n$  es número de compartimentos del histograma que se activaron por lo menos una vez con alguna superficie en la base de datos.

### 3.2 Generación de histogramas a partir de imágenes reales

Para obtener el mismo tipo de histogramas tridimensionales descritos en la sección inmediatamente

anterior, a partir de imágenes reales, se siguió con algunas variaciones el procedimiento descrito por Barnard *et al.* (2002), seguido también por Funt y Xiong (2004) entre otros.

Se usaron las 321 imágenes reales del set recopilado por Barnard *et al.* (2002a). A cada una de estas imágenes se le aplicó un filtro promedio de  $5 \times 5$ . Luego se descartaron aquellos píxeles con valor de intensidad menor que 8 (en una escala de 0 a 255) en alguno de sus canales (Cardei, 2000). Con los píxeles resultantes se calculó el histograma tridimensional en el espacio espectral elegido, usando una distribución de retícula cartesiana, como se explicó para las imágenes sintéticas.

Para los experimentos con imágenes reales, el proceso de poda se limitó a descartar los nodos que resultaron inactivos en los experimentos correspondientes con imágenes sintéticas, es decir, aquellos que poseían los mismos parámetros (espacio espectral de color y tamaño de arista).

### 3.3 Implementaciones de la regresión contraída y del perceptrón multicapa

Para implementar la regresión contraída se agruparon convenientemente los datos de la manera que lo muestran las ecuaciones (4) y (5). Se calcularon los coeficientes de la regresión como lo muestra la ecuación (11).

Para aprovechar el parámetro  $\lambda$ , se calcularon diferentes coeficientes para diferentes valores de  $\lambda$  y se graficaron, mediante la ecuación (12), los errores obtenidos en cada caso contra el parámetro  $\lambda$ , eligiendo el del menor error, como el del  $\lambda$  óptimo.

Dada la inestabilidad de los resultados obtenidos con las redes neuronales artificiales ante un

mismo set de entrenamiento, se adoptó una estrategia en la que se entrenaba una batería de 48 perceptrones multicapa con el mismo número de capas, nodos por capa y las mismas funciones de activación, para observar los resultados que se obtenían.

En la sección de resultados del presente trabajo se describirán las arquitecturas neuronales implementadas para cada experimento. El rasgo común de dichas arquitecturas es que se trata de perceptrones multicapa, con funciones de activación logaritmo sigmoideal, entrenados por retro-propagación del error, con tasa de aprendizaje elástica y parada temprana con un set de entrenamiento y validación de 17.400 imágenes sintéticas cada uno.

Lo que varía para los diferentes experimentos es el número de neuronas en cada capa, la generación de las imágenes sintéticas, los espacios espectrales utilizados y la prueba que puede hacerse con imágenes sintéticas o reales.

Las implementaciones de la regresión contraída y el perceptrón multicapa se llevaron a cabo en Matlab.

### 3.4 El enfoque cooperativo entre estas dos máquinas de aprendizaje

En los antecedentes se ha mostrado la regresión contraída como una técnica lineal con parámetro de regularización. Por ser lineal, su implementación a través de matrices es más fácil y requiere menos recursos que otras máquinas de aprendizaje (máquinas de soporte vectorial y redes neuronales según Argawal (2005)). Gracias a su parámetro de regularización, la regresión contraída ofrece mejores resultados que una regresión lineal simple.

De otro lado, aunque la inestabilidad de las redes neuronales se ha

demostrado en trabajos anteriores (Argawal, 2005; Gómez, Loaiza y Caicedo, 2006b), se ha comentado acerca de la gran cantidad de parámetros por optimizar (Argawal, 2005) y se ha explorado a través de pruebas preliminares que la complejidad computacional exige grandes tiempos computacionales.

Sin embargo, las comparaciones iniciales entre estas dos máquinas de aprendizaje, llevadas a cabo por los autores del presente artículo, han mostrado algunos resultados con redes neuronales, superiores a los obtenidos con regresión contraída.

En este orden de ideas, la propuesta central de este artículo es el aprovechamiento de la regresión contraída como técnica exploratoria, por su facilidad de implementación, alta velocidad de cómputo y estabilidad, luego con los resultados de la exploración se lleva a cabo un entrenamiento con una batería de redes neuronales artificiales, de las cuales se espera un mejor desempeño.

Este enfoque cooperativo y no competitivo es original de los autores y representa un ahorro en tiempo y recursos computacionales. Para la exploración de parámetros en la estimación de iluminantes y los ejemplos de este enfoque, al igual que mayores detalles acerca del proceso seguido, pueden constatarse en la sección de resultados.

### 3.5 Medidas de error

Se consideraron dos errores clásicos en estimación del iluminante: la distancia euclidiana y el error angular.

El primer error, como su nombre lo indica, es la distancia euclidiana calculada en el plano de cromaticidad  $(r, g)$ , entre el iluminante estimado  $(r_e, g_e)$  por la máquina de aprendizaje y el iluminante verdadero – “target” en la literatura en inglés –  $(r_t, g_t)$ , medido en la base

de imágenes reales (Barnard *et al.*, 2002 y 2002a) o conocido en la generación de imágenes sintéticas.

$$\text{Dist.Euclidiana} = \sqrt{(r_i - r_e)^2 + (g_i - g_e)^2} \quad (15)$$

Siguiendo a Cardei (2000) se usó el promedio de las distancias euclidianas para un grupo de imágenes para dar una sola medida por cada repetición.

El segundo error, es el ángulo existente entre los vectores tridimensionales que representan al iluminante estimado ( $r_e, g_e, b_e$ ) y al iluminante verdadero ( $r_i, g_i, b_i$ ).

$$\text{ErrorAng.} = \frac{\pi}{180} \cos^{-1} \left( \frac{(r_i, g_i, b_i) \bullet (r_e, g_e, b_e)}{\sqrt{r_i^2 + g_i^2 + b_i^2} \sqrt{r_e^2 + g_e^2 + b_e^2}} \right) \quad (16)$$

Para calcular la componente  $b$  de cada vector tridimensional bastó recordar que:

$$b = 1 - \sqrt{r^2 + g^2} \quad (17)$$

Al igual como lo hicieron Barnard, Cardei y Funt (2002) se usó el valor cuadrático medio (RMS) de los errores angulares de un grupo de imágenes para comparar los resultados.

$$\text{RMSErrorAng} = \frac{(\text{ErrorAng.})}{\sqrt{N}} \quad (17)$$

Donde (ErrorAng.) es el vector con todos los errores angulares (uno por imagen) calculados para una repetición, de dimensiones  $[1 \times N]$ , es decir de  $N$  imágenes.

#### 4. Resultados

Aprovechando el enfoque cooperativo propuesto, se exploró el efecto del tamaño de la arista en los espacios espectrales RGBdispositivo y rgL, mediante la regresión contraída (Secciones 0 y 0) para luego implementar RNA (Sección 0) con mejores resultados que los obtenidos hasta ahora por los autores.

De igual manera se exploró la normalización del set de iluminantes con la regresión contraída y se eligió la mejor para implementar 48 MLP (Sección 0); a continuación se presentan los resultados, que muestran cómo la herramienta permite gran cantidad de experimentación exploratoria antes de invertir tiempo en el entrenamiento de una batería de RNA.

##### 4.1 Exploración inicial del tamaño de la arista

Pruebas preliminares a estos experimentos plantearon la posibilidad de usar la regresión contraída para explorar el efecto de la disminución del tamaño de la arista del histograma tridimensional sobre el desempeño del algoritmo, invirtiendo menos tiempo y con mayor estabilidad que los perceptrones multicapa.

Para ello se entrenaron y optimizaron por validación cruzada a partir de imágenes sintéticas (set inicial), un grupo de coeficientes de regresión contraída para distintos valores de arista,  $a$  y se consignaron sus resultados en la Tabla 1.

$a$	Distancia Euclidiana	Error Angular
0.5	0.0576	9.3020°
0.2	0.0376	6.2485°
0.1	0.0343	4.5938°
0.08	0.0342	5.7246°
0.06	0.0335	5.6393°
0.044	0.0337	5.7059°

**Tabla 1.** Errores de prueba para distintas aristas con regresión contraída ante el set inicial en el espacio RGB dispositivo.

Analizando la Tabla 1 se puede apreciar que a medida que disminuye el tamaño de la arista, la distancia euclidiana se hace más pequeña, pero que el comportamiento del error angular no presenta exactamente la misma tendencia.

También llamó la atención que el valor más pequeño de arista



(0.044) no presentara los errores más bajos, pero inicialmente solo se observó como una curiosidad.

#### 4.2 Implementación del espacio espectral rgL

Dado que Funt y Xiong (2004) y Barnard, Martin y Funt (2000) habían sido exitosos en el tránsito de 2D a 3D usando el espacio espectral rgL, se aprovechó la herramienta de la regresión contraída para probar rápidamente una propuesta de implementación en este espacio espectral.

De otro lado, las pruebas preliminares permitieron observar que los errores de validación eran inferiores a los de prueba, así que se tomó la característica de cota para afirmar que si se usaba el set de prueba, los errores serían superiores a los obtenidos con el set de validación, así que no valía la pena someter a prueba aquellos casos en los cuales los errores en la validación fuesen altos.

El otro aspecto que se aprovechó para probarse en este ensayo, fue la introducción de un nuevo set de imágenes sintéticas que se generó evitando la normalización de los espectros de los iluminantes.

a	Distancia Euclidiana	Error Angular
0.5	0.0649	10.0416°
0.2	0.0354	5.5125°
0.1	0.0298	4.9623°
0.08	0.0308	5.1499°
0.06	0.0311	5.2927°
0.044	0.0321	5.5600°

**Tabla 2.** Errores de validación para distintas aristas con regresión contraída ante el set sin normalización de iluminantes en el espacio rgL.

Al comparar las distancias euclidianas y los errores angulares de la Tabla 2 con las de la Tabla 1, se puede apreciar que salvo para la arista de 0.5, los errores son inferiores en el espacio rgL.

En la Tabla 2 se puede observar que el menor error (para ambos errores) se presentó con la arista de 0.1.

De igual manera, basados en los resultados, se prefirió seguir usando el set de datos sin normalización de iluminantes en las siguientes pruebas.

Y para corroborar su ventaja se implementó una prueba en el espacio RGB dispositivo, pero con el set sin normalización de los iluminantes.

a	Distancia Euclidiana	Error Angular
0.5	0.0576	9.3020°
0.2	0.0376	6.2485°
0.1	0.0343	4.5938°

**Tabla 3.** Errores de validación para distintas aristas con regresión contraída ante el set sin normalización de iluminantes en el espacio RGB dispositivo.

Si se comparan las distancias euclidianas de la Tabla 3 con las primeras tres filas de la Tabla 1, se puede apreciar la ventaja de usar el set sin normalización de iluminantes.

A continuación se propuso comparar el enfoque bidimensional rg con el tridimensional rgL, ya que esta comparación fue una de las motivaciones de la tesis.

Para hacerlo se implementó una regresión contraída en el plano rg con una arista de 0.02 y se comparó con su contraparte tridimensional (rgL) conservando el tamaño de la arista para el plano rg (0.02) y con el máximo número de elementos que soportaba la memoria de Matlab, que se logró con una arista independiente para la luminancia L de 0.25.

A la anterior comparación se adicionó una implementación en el plano rg, con el mínimo tamaño de arista soportado por Matlab en red

neuronal, pues aun cuando se puede implementar un tamaño de arista inferior con regresión contraída no se debe perder de vista el objetivo de usarla como cota para RNA (ya que al implementarla en Matlab, no puede reducirse mucho el tamaño de la arista, pues se aumentaría el número de nodos de entrada colapsando la memoria).

Resulta interesante que al incorporar la tercera dimensión L, no necesariamente mejora el desempeño (comparar filas 1 y 2 de la Tabla 4) y de nuevo se observa que el menor tamaño de arista tampoco garantiza mejor desempeño (comparar filas 1 y 3 de la Tabla 4).

Espacio	a	SL	Distancia euclidiana	Error angular
rg	0.02	n.a.	0.0318	5.41°
rgL	0.02	0.25	0.0323	5.57°
rg	0.018	n.a.	0.0326	5.59°

**Tabla 4.** Comparación de implementaciones rg y rgL.

Aprovechando que se había modificado el programa de regresión contraída para poner valores distintos al plano rg y la dimensión L, se implementó un SL de 0.1 y se varió la arista del plano rg.

En la Tabla 5 se puede ver que los errores vuelven a tener un mínimo en la arista de 0.1 y si se comparan los errores con los de la Tabla 2, se nota que el valor óptimo de arista (aquel con el menor error) no es el menor valor de arista implementado.

a	Distancia euclidiana	Error angular
0.5	0.0464	7.4032°
0.2	0.0324	5.1681°
0.1	0.0298	4.9623°
0.08	0.0302	5.0431°
0.06	0.0304	5.1395°

**Tabla 5.** Errores de validación para distintas aristas en rgL con SL=0.1.

Dado que el mejor valor hasta el momento se obtuvo con una arista de 0.1 en el plano rgL y que dicho valor se logró sobre el set sin normalización de iluminantes, se quiso verificar su desempeño en el set con normalización de iluminantes.

Normalización de iluminantes	Distancia euclidiana	Error angular
SI	0.0298	5.0068°
NO	0.0298	4.9623°

**Tabla 6.** Efecto de la normalización de los iluminantes con una arista de 0.1 en el espacio rgL.

Como se puede ver en la Tabla 6, el efecto de la normalización de los iluminantes no es clave para este tamaño de arista y se optó por seguir usando en las pruebas el set sin normalización de los iluminantes, por sus buenos resultados en otros tamaños de arista.

Manteniendo el concepto de cota exploratoria planteado en esta aproximación, se compararon los valores promedio de distancia euclidiana, obtenidos con las RNA implementadas en los meses anteriores y las obtenidas con las regresiones contraídas.

Espacio	a	RNA	RR	Diferencia
rg	0.018	0.0256	0.0326	0.0070
RGB dispositivo	0.08	0.0220	0.0342	0.0122
RGB dispositivo	0.06	0.0255	0.0335	0.0080

**Tabla 7.** Comparación de las distancias euclidianas obtenidas con RNA y regresión contraída.

De acuerdo con la columna 5 de la Tabla 7, se esperaba una mejora en la implementación mediante RNA de por lo menos 0.070 y se eligió la arista de 0.1 en el espacio rgL, ya que las pruebas con regresión contraída la mostraron como la mejor candidata.

#### 4.3 Implementación RNA con los parámetros elegidos gracias a la exploración mediante regresión contraída

Usando una arista de 0.1 en el espacio espectral rgL, con imágenes sintéticas generadas sin normalización del iluminante, tal y como se había establecido con la exploración mediante regresión contraída de la sección anterior, se procedió a entrenar, validar y probar 48 MLP con función de activación logarítmico sigmoideal y con el siguiente número de neuronas por capa [208 50 20 2].

Se usó entrenamiento por retro-propagación del error con tasa de aprendizaje elástica y parada temprana, ante set de entrenamiento, validación y prueba de 17.400 imágenes sintéticas cada uno.

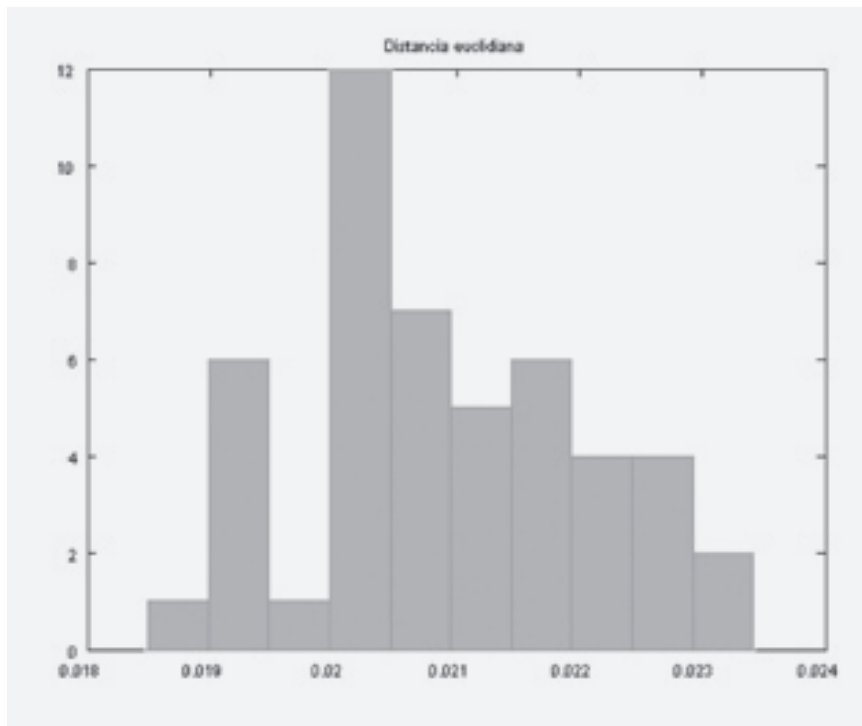


Figura 2. Histograma de las distancias euclidianas promedio de 48 MLP's con una arista de 0.1 en el espacio rgL.

En la Figura 2 puede observarse que se obtuvieron distancias euclidianas que no superan las 0.0240 unidades adimensionales, con un desempeño aceptable para una red

con menos neuronas que cualquiera de las redes implementadas en el plano rg en la tesis de Cardei (2000).

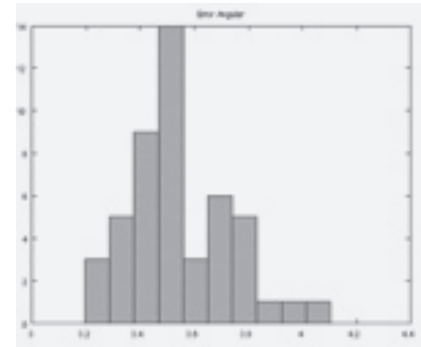


Figura 3. Histograma de los errores angulares (RMS) de 48 MLP's con una arista de 0.1 en el espacio rgL.

Los resultados de los errores angulares (Figura 3), no se pueden comparar con los de Cardei (2000), pues en su tesis no se usó el error angular como medida de error, pero resultan aceptables si se comparan (Tabla 8) con la mejor red que se ha implementado hasta el momento por Gómez, Loaiza y Caicedo (2006b).

Espacio Espectral	a	Arquitectura	Distancia euclidiana	Error angular
RGB disposit.	0,08	[524 300 150 2]	0,0220	3,8826°
rgL	0,10	[208 50 20 2]	0,0209	3,5499°

Tabla 8. Comparación de las RNA implementadas.

En la Tabla 8 puede apreciarse la ventaja de este enfoque para explorar el posible desempeño de una RNA respecto a los resultados obtenidos con lentas exploraciones usando redes neuronales artificiales.

En la Figura 4, puede verse la razón por la cual resultó efectivo, porque para estos datos la regresión contraída se comportó como una

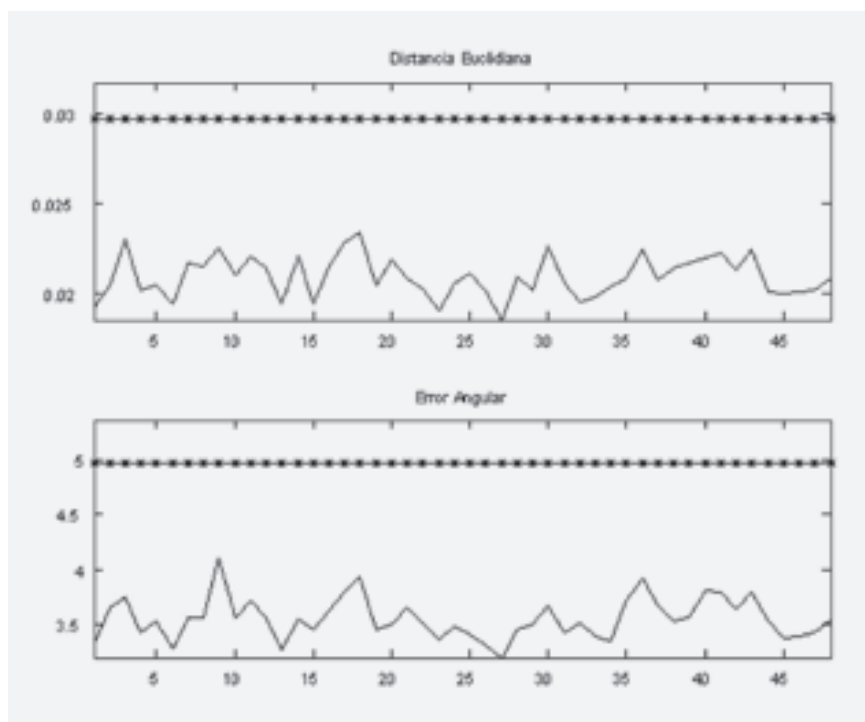


Figura 4. La regresión contraída (-\*) como cota para la implementación de 48 MLP (-).

cota del desempeño de los MLP implementados con los mismos datos.

El paso siguiente fue explorar una red más compleja. Se entrenó entonces la batería de 48 MLP pero usando una arquitectura un poco más compleja, en cuanto al número de neuronas [208 100 50 2].

Los resultados no fueron tan buenos como se esperaban pues el desempeño fue comparable con el del MLP más sencillo (Ver segunda fila de la Tabla 8): la misma distancia euclidiana promedio (0.0209) y un error angular promedio de 3.5384°.

Esto representa un posible aporte al trabajo de Cardei (2000) y coincide con algo que se había observado en el trabajo de Argawal (2005): las redes quizás no necesitaban tantas neuronas para alcanzar un buen funcionamiento ya que en Argawal, 2005 se usaron menos neuronas que en Cardei, 2000, e incluso una sola capa oculta, aunque el entrenamiento en Argawal, 2005 se hizo directamente con imágenes reales.

Se hizo otra comparación usando los valores mínimos y máximos, para

la distancia euclidiana (DE) y el Error Angular (EA) y aunque se aprecian pequeñas diferencias, siguen siendo comparables los resultados.

#### 4.4 Exploración de normalizaciones para mejorar el paso de imágenes sintéticas a reales en el plano rg.

Durante las pruebas preliminares se observó que algunas implementaciones que funcionaban muy bien con imágenes sintéticas, no ofrecían buenos resultados con imágenes reales y una de las posibles causas era una normalización errada en el modelo de generación de imágenes sintéticas.

Hasta este punto de la experimentación se había usado un set de imágenes sintéticas generado con

		Min	Media	Max
DE	Compleja	0.0178	0.0209	0.0241
	Simple	0.0185	0.0209	0.0234
EA	Compleja	3.0918	3.5384	4.1005
	Simple	3.1955	3.5499	4.1046

Tabla 9. Comparación de dos arquitecturas Neuronales.

una normalización de los iluminantes, dividiendo cada espectro entre la sumatoria de los valores para todas sus longitudes de onda para comparar las formas de los espectros y no su magnitud (etiquetada como “Óptica” en la Tabla 10).

Considerando como posible causa del fracaso en el paso de imágenes sintéticas a reales, la normalización llevada a cabo en la sección anterior (ver como ejemplos Tabla 2 y Tabla 6), se decidió probar el uso de un set de imágenes generadas sin normalizar los iluminantes (Etiquetada como “Sin” en la Tabla 10).

Para completar esta exploración se probó normalizar los iluminantes de la forma descrita en la página 974 de Barnard *et al.*, 2002, para que ante un blanco perfecto la máxima salida de entre los tres canales fuera 255.

En la mayoría de los experimentos el Set de imágenes se ha sintetizado a partir de 251 reflectancias, para tratar de seguir lo descrito en Cardei, 2000 y Cardei, Funt y Barnard, 2002, pero en esta oportunidad se aprovechó para implementar usando las 1.995 superficies d (Barnard *et al.*, 2002, como se hizo en Funt y Xiong, 2004 y Barnard *et al.*, 2002, para establecer futuras comparaciones.

La prueba comparativa de estas cuatro opciones se llevó a cabo en el plano rg (enfoque bidimensional) con una arista de 0.018 (la mínima que soportó Matlab para entrenar RNA) y se hizo usando regresión contraída por las ventajas ya expuestas de esta técnica.

Tipo de Normalización	Distancia Euclidiana	Error Angular
Óptica	0.0323	5.5273
255	0.0325	5.5549
Sin	0.0321	5.5079
255 para 1995 superficies	0.0352	5.8427

**Tabla 10.** Comparación de los tipos de normalización

De estos resultados (Tabla 10), llamó la atención que el mejor desempeño se obtuvo cuando no se normalizaban los iluminantes y el peor, si se usaban las 1.995 superficies.

Se observó que cuando se usaron 251 reflectancias para generar los datos sintético, la primera capa podada quedaba con 1.025 nodos que se activaban, mientras al usar 1995 reflectancias el número de nodos activos en la primera capa era de 1.123.

Con estos resultados de las regresiones contraídas se pasó a las implementaciones con RNA, usando la siguiente estructura: una batería de 48 MLP con función de activación logaritmo sigmoideal [1025 50 20 2] neuronas por capa, se usó entrenamiento por retro-propagación del error, con tasa de aprendizaje elástica y parada temprana, set de entrenamiento y validación de 17.400 imágenes sintéticas cada uno y un set de prueba con las imágenes reales descritas en la sección 0.

	Min.	Media	Máx.
Distancia Euclidiana	0.0589	0.0835	0.1233
Error Angular	8.8584	12.4789	21.6222

**Tabla 11.** Errores al entrenar 48 MLP con imágenes sintéticas generadas con normalización óptica de los iluminantes y probar con imágenes reales en el plano rg

Esta vez los errores mínimos resultaron similares a los obtenidos en Cardei, 2000 y Barnard *et al.*, 2002 y esta coincidencia con la literatura verificó que una RNA operando en el plano rg, puede entrenarse con imágenes sintéticas y tener un desempeño aceptable con imágenes reales. Aunque se debe destacar que sólo algunas redes de este trabajo obtuvieron datos coincidentes con los de la literatura.

Aplicando el mismo tipo de prueba, pero para el set de imágenes sintéticas generadas sin normalización de los iluminantes, ofreció los siguientes resultados.



	Min.	Media	Máx.
Distancia Euclidiana	0.0545	0.0758	0.1303
Error Angular	8.2356	11.4385	18.7983

**Tabla 12.** Errores al entrenar 48 MLP con imágenes sintéticas generadas sin normalización de los iluminantes y probar con imágenes reales en el plano rg

Como se puede observar, al comparar la Tabla 11 con la Tabla 12, el set de imágenes sintéticas sin normalización de iluminantes, tuvo mejores resultados aun cuando se probara con imágenes reales.

## 5. Conclusiones

El aporte principal de este trabajo es la propuesta de un enfoque cooperativo entre dos máquinas de aprendizaje: perceptrones multicapa y regresión contraída.

La ventaja de este enfoque es el ahorro de tiempo de experimentación y de cómputo resultante de aprovechar la facilidad de implementación y velocidad de cómputo de la regresión contraída en una exploración de parámetro, antes de lanzar una prueba más dispendiosa con redes neuronales artificiales.

El caso particular de estudio: estimación de iluminantes, el enfoque resultó exitoso, llevando a la implementación de redes neuronales de mejor desempeño, con menor esfuerzo computacional y menos tiempo invertido en experimentación, que en trabajos anteriores (Gómez, Loaiza y Caicedo, 2006b).

Específicamente en la estimación de iluminantes en imágenes sintéticas, se pudo establecer que el tamaño óptimo de la arista de los histogramas tridimensionales no fue el menor tamaño de arista implementado.

Asimismo, aunque los mejores resultados en este trabajo se obtuvieron entrenando con imágenes sintéticas sin normalización de iluminantes y partiendo de 251 reflectancias, se recomienda al investigador inte-

resado normalizar a 255 y usar las 1.995 reflectancias disponibles para ceñirse a la literatura.

En cuanto a las direcciones futuras, los autores se encuentran trabajando en la definición de un protocolo para la evaluación de algoritmos de estimación de iluminantes, que facilite la comparación del enfoque cooperativo con otros métodos de estimación del iluminante clásico y se espera llevar a cabo una aplicación demostrativa para el algoritmo que se ha propuesto.

## 6. Agradecimientos

Los autores agradecen a Colciencias por patrocinar la realización de este artículo, dentro del marco del programa: “Apoyo a doctorados nacionales”, concedido a Edwin Octavio Gómez González; también agradecen al programa de doctorado en Ingeniería de la Universidad del Valle. ☼

## 7. Bibliografía

- Argawal, V. 2005. Ridge Regression approach to Color Constancy, M. Sc. dissertation Knoxville, TN, USA: the University of Tennessee., [Online]. Disponible en línea en: [http://imaging.utk.edu/people/former/vivek/Thesis\\_VivekAgarwal.pdf](http://imaging.utk.edu/people/former/vivek/Thesis_VivekAgarwal.pdf)
- Barnard, K.; Cardel, V.; Funt, B. 2002. A Comparison of Computational Color Constancy Algorithms; Part One: Methodology and Experiments with Synthesized Data. *IEEE Transactions on Image Processing*, 11(9). Pp. 972-984.
- Barnard, K.; Martin, L.; Coath, A.; Funt, B. 2002. A Comparison of Computational Color Constancy Algorithms; Part Two: Experiments with Image Data. *IEEE Transactions on Image Processing*, 11(9). Pp. 985-996.
- Barnard, K.; Martin, L.; Funt, B.; Coath, A. 2002a. A Data Set for Colour Research, *Color Research and Application*, Volume 27, Number 3, pp. 147-151
- Barnard, K; Martin, L; funt, B. 2000. Colour by correlation in a three-dimensional colour space. *ECCV'2000*
- Proceedings Sixth European Conference on Computer Vision, Dublin, 2000. pp 275-289.
- Buchsbaum, G. 1980. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310. Pp. 1-26.
- Cardei, V. 2000. A Neural Network Approach to Color Constancy. Ph.D. dissertation, Simon Fraser University School of Computer Science. Disponible en línea en: [http://www.cs.sfu.ca/%7Ecolour/publications/VCardeiPhD/VCardei\\_PhD\\_Thesis.pdf](http://www.cs.sfu.ca/%7Ecolour/publications/VCardeiPhD/VCardei_PhD_Thesis.pdf) Burnaby, BC, Canadá. Dirección verificada 13 de Julio de 2006.
- Finlayson, G.D.; Trezzi, E. Shades of gray and colour constancy. In *IS&T/SID Twelfth Color Imaging Conference*, pages 37-41, 2004.
- Funt, B.; Cardel, V.; Barnard, K. 1996. Learning Color Constancy. In *Proceedings of the Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*. Scottsdale, AZ. November, 1995. Volume 4. Pp. 58-60.
- Funt, B.; Xiong, W. 2004. Estimating Illumination Chromaticity Via Support Vector Regression. In *Proceedings of the Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*. Scottsdale, AZ. November 9, 2004. Volume 12. Pp. 29-36.
- Gómez, E.; Loaiza, H.; Caicedo, E. 2006a. Taxonomía de la constancia del color computacional, una aproximación. *El hombre y la máquina*, 25. Pp. 30 - 43.
- Gómez, E.; Loaiza, H.; Caicedo, E. 2006b. Estimación del Iluminante en Imágenes Sintéticas Simples mediante Redes Neuronales Artificiales operando sobre Espacios Espectrales Tridimensionales. *XI Simposio de Tratamiento de Señales, Imágenes y visión artificial, STSIVA 2006*.
- Rosenberg, C.; Hebert, M.; Thrun, S. 2001. Color constancy using KL-divergence. In *Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV*, 7-14 July 2001. Volume 1. Pp. 239-246.
- Van de Weijer, J; Gevers, TH. 2005. Color Constancy based on the gray-edge hypothesis. *IEEE International Conference on Image Processing, ICIP 2005*. Volume 2, 11- 14 Sept. 2005 Page(s): II- 722-5